形式言語の複雑さに関する研究

著者	五十嵐 善英
号	280
発行年	1970
URL	http://hdl.handle.net/10097/9016

氏 名（本籍）	五十嵐 善 英（宮城県）
学位の種類	工 学 博 士
学位記番号	工 博 第 280 号
学位授与年月日	昭和46年3月25日
学位授与の要件	学位規則第5条第1項該当
研究科専門課程	東北大学大学院工学研究科 （博士課程）電気及通信工学専攻
学位論文題目	形式言語の複雑さに関する研究
論文審査委員	（主査） 教授 本多 波雄 教授 大泉 充郎 教授 城戸 健一 教授 木村 正行

論 文 内 容 要 旨

ま え が き

句構造言語は，その生成文法の形やそれを受理するオートマトンの形により，context－sensitive言語，決定性context－sensitive 言語，非決定性1方向スタック言語，context－free 言語，決定性context－free 言語，linear context－free言語，正規集合等に分類されている。このような分類と違って，長さnの語が標準のチューリング機械によってT（n）ステップ以下の操作で認識されるとき，T（n）一認識可能とし，それによって句構造言語の複雑さのクラスを定義したり，同様な方法で，認識に必要なテープの量から複雑さのクラスを定義することが試みられている。

言語の認識の複雑さについては多くの研究がなされているが，言語の生成の過程の基本的な性質

を理解し，生成の複雑さに基ずく言語のクラスを研究することも重要である。context－sen-sitive 文法の導出の複雑さについては，A．V．Gladkij がはじめに研究し，興味ある結果を示した。R．V．Bookは，A．V．Gladkij の導出の複雑さを句構造文法に拡張し，いくつかの結果を導いた。本論文ではA．V．Gladkij やR．V．Book 等の定義と異る導出の複雑さの測度を与え，それによって定義される言語の複雑さのクラスに関する結果を導く。

第１章　導出の複雑さ

この章では，導出の複雑を定義し一般的な性質を述べる。

定義域が全非負整数であり，すべての非負整数nについて，$T(n+1) \geq T(n)$ なる非負整数から正整数への計算可能な関数を時間関数という。

Tを時間関数，αを句構造言語とする。各非負整数nについて，αの長さnのすべての語を$T(n)$以下の長さの導出で生成できる句構造文法が存在するときαをT－導出言語という。すべてのT－導出言語のクラスをL_Tで表わす。

任意の時間関数TについてL_Tは帰納的に可算な集合であり，L_Tを真に含むような時間関数Uが存在することが示される。したがって，異ったクラスの無限に長い鎖，$L_{T1} \subsetneqq L_{T_2} \subsetneqq \cdots\cdots$が存在する。

A．V．Gladkijがcontext-sensitive文法上で定義したconnected文法を句構造文法で定義し，これに関する結果を与える。connected文法に関する結果からαがT－導出言語であることが導かれる。

導出の複雑さの測度を用いて，非負整数から非負整数への計算可能な関数の複雑さを定義することができる。fをすべての非負整数から非負整数への計算可能な関数とし，cを{0，1}に含まれない記号とする。$L(G)=\{1^n c 0^{f(n)} \mid n \geq 0\}$でかつ，各nについて，$1^n c 0^{f(n)}$を生成する導出の長さが$T(n)$を越えないような句構造文法Gが存在するときfをT－構成可能関数という。

第２章　クラスL_T上の演算

L_Tについて，閉じている演算と閉じていない演算を述べる。主な結果は以下のとおりである。

任意の時間関数TについてL_Tはε－freeな準同型写像，reverse，和集合，積集合，正規集合との共通集合の各演算で閉じている。Tをかってな正整数kについて，$\inf_{n\to\infty}(T(n)/T(kn)) \gneqq 0$ なる時間関数とすると，L_Tは逆準同型写像の演算で閉じている。Tを$\inf_{n\to\infty}(T(n)/n^2) \gneqq 0$ なる時間関数とすると，L_Tは共通集合の演算で閉じている。Tをすべて

の正整数の組n_1, n_2について，$T(n_1+n_2) \geq T(n_1)+T(n_2)$なる時間関数とすると，$L_T$は*演算で閉じている。Tを$\underset{n\to\infty}{\mathrm{Inf}}(T(n)/n) \gneq 0$なる時間関数とすると，$L_T$は準同型写像，正規集合によるright quotient，正規集合によるleft quotientの各演算で閉じていない。Tを$\underset{n\to\infty}{\mathrm{Inf}}(T(n)/n) \gneq 0$でかつ$\lim_{n\to\infty}(T(n)/n^2)=0$なる時間関数とすると，$L_T$は共通集合，補集合の各演算で閉じていない。

第3章　言語クラス間の関係

この章では，各言語クラスと複雑さのクラスL_Tとの関係について述べる。context－free言語のクラスはL_nに真に含まれることがまず示される。S．GinsburgとS．A．Greibachが研究したabstract family of languages（AFLと略記する）とfull abstract family of languages（full AFLと略記する）についてはつぎのことが示される。任意の時間関数Tについて，L_Tはfull AFLではない。つぎの（i）と（ii）の条件を満たす時間関数Tについて，L_TはAFLである；（i）$\underset{n\to\infty}{\mathrm{Inf}}(T(n)/T(kn)) \gneq 0$ ただしkはかってな正整数，（ii）すべての正整数の組n_1，n_2について$T(n_1+n_2) \geq T(n_1)+T(n_2)$。よって，任意の正整数rについて，$n^r$－導出言語のクラスはAFLである。R．V．BookとS．A．Greibackによって研究された準実時間言語のクラスはL_nを真に含み，L_{n^2}に含まれることが示される。ε－freeなn－導出言語のクラスは決定性context－sentive言語のクラスに真に含まれ，$L=\cup_{k=1}^{\infty}(L_{k^n})$は，context－sensitive言語のクラスを含むことが示される。ついで，S．GinsburgとS．A．Greibackによって研究されたabstract family of acceptors（AFAと略記する），abstract family of transducers（AFTと略記する）との関係が述べられる。これらの結果からTを，$\underset{n\to\infty}{\mathrm{Inf}}(T(n)/n) \gneq 0$なる時間関数とすると$L_T$は1方向非決定性スタック言語のクラスに含まれないことが示される。

第4章　ハイアラーキ

まず，L_nより小さいクラスについて論じる。$\underset{n\to\infty}{\mathrm{Sup}}(T(n)/n) \lneq \infty$でかつ，かってな正整数kについて，$T(n) \geq n/k$を満さない非負整数nが存在する時間関数を$T \ll n$と書き，かってな正整数kについて$T(n) \geq n/k$を満す非負整数nが有限個しか存在しない時間関数を，$T \lll n$と書く。$T \lll n$なる任意の時間関数Tについて，L_Tはすべての有限集合のクラスである。$T \ll n$で，かつL_Tが無限集合の言語を含むようなTは存在する。かってな正整数kについて，$T(n) \geq n/k$を満さない非負整数nが存在するような時間関数Tについて，L_Tは{｜w｜

$|w\in\alpha\}$がultimately periodicな無限集合になるαを含まない。ただし，$|w|$は語wの長さを表わす。$T\ll n$なる任意の時間関数Tについて，L_TはL_nに真に含まれる。以上のような結果が，まず導かれる。ついでAFLとの関係を述べ，AFL を含む無限に長いハイアラーキの鎖が存在することが示される。すなわち，任意のL_Tについて$L_T \subsetneq F \subsetneq L_Q$なる$L_Q$とAFL，F が存在する。したがって，$L_{T1} \subsetneq F_1 \subsetneq L_{T2} \subsetneq F_2 \subsetneq \cdots\cdots$ なる無限に長い鎖が存在する。ここで，各F_i（$i=1，2\cdots\cdots$ ）はAFLである。ついで，A.v.Gladkijが導いたcontext－sensitive言語に関する非線形定理を拡張したより一般的な定理を与える。φを，$\Sigma=\{a_1, a_2\}$上のすべての系列の上で定義されている関数で，φの関数値はΣ^*の空でない部分集合とし，$\mathcal{L}_\varphi=\{xb\varphi(x)bx \mid x\in\Sigma^*\}$ とする。ただし，bはΣに含まれない記号とする。$F\varphi$を，Σ^*から正整数への関数で，$\mathcal{F}\varphi(x)=\min\{|xb\varphi(x)bx|\}$とし，$f\varphi(n)=\max\{\mathcal{F}\varphi(x) \mid |x|=n\}$とする。Tを，$\lim_{n\to\infty}(T(f\varphi(n))/n^2)=0$ なる時間関数とするとL_Tは$\mathcal{L}_\varphi$を含まないことが導びかれる。この結果は，L_nとL_{n^2}の間のハイアラーキの問題を解くのに有用である。$\beta_1=\{xba_1^{|x|[\sqrt{|x|}]}bx \mid x\in\Sigma^*\}$とすると，$\beta_1$は$n^{\frac{4}{3}}$導出言語であるが，上の結果を用いて$\lim_{n\to\infty}(T(n)/n^{\frac{4}{3}})=0$なる時間関数Tでは，$L_T$は$\beta_1$を含まない（$[r]$は，$m\geqq r$なる最小の正整数mを表わす）ことが示される。同様な方法で，$\beta_2=\{xba_1^{|x|[\sqrt[3]{|x|}]}bx \mid x\in\Sigma^*\}$は$n^{\frac{2}{3}}$－導出言語であるが，$\lim_{n\to\infty}(T(n)/n^{\frac{3}{2}})=0$なる時間関数Tについて$L_T$は$\beta_2$を含まないことが示される。同様な方法を続けることによって，つぎの結果がえられる。L_nとL_{n^2}の間に真の包含関係がある無限に長い鎖，$L_n \subsetneq L_{T1} \subsetneq L_{T2} \subsetneq \cdots\cdots \subsetneq L_{n^2}$が存在する。また$\gamma=\{xba_1^{|x|[\log_2[x]]}bx \mid x\in\Sigma^*\}$とすると，$k \lneq 2$ なるどのような計算可能な正数kについても，n^k－導出言語でないことが示される。このことから$(n/\log n)^2$ -導出言語のクラスはn^2 -導出言語のクラスに真に含まれ$k \lneq 2$なる任意の計算可能な正数kについてn^k -導出言語のクラスを真に含むことが示される。（対数の底は1よりも大の任意の数である）。

この章の終りの節で，2つのクラスL_TとL_QについてL_TがL_Qを真に含むための十分条件を与える。すなわち，つぎの条件を満すT-構成可能関数fが存在するならば，L_TはL_Qを真に含む；$\operatorname{Inf}_{n\to\infty}(T(n)/(2^{f(n)}f(n)Q(n)\log n)) \gneq 0$でかつ，$\operatorname{Inf}_{n\to\infty}(f(n)/Q(n)\log(\log n)) \gneq 0$（対数の底は1より大の任意の数）。この結果を用いて，context－sensitive言語のクラスは，$2^{[\log_p{}^{k}[n]]^{rn}}$-導出言語のクラスに真に含まれることが示される。ここで，kは任意の正整数，rは任意の計算可能な正数，pは1より大の計算可能な数，$\log^1[n]=\log[n]$，$\log^{r+1}[n]=\log[\log^r[n]]$とする。

第５章　決定不能問題

この章では，クラスL_Tに関する決定問題について述べる。context-free言語に関する決定問題や，S.A.Greibach による一般的に取り扱った決定問題の結果から，多くの決定問題を解くことができる。

あ と が き

本論文で議論されている複雑さのクラスのハイアラーキについては，とくに興味があり，未解決な問題も多い。本論文ではL_nとL_{n2}の間に真の包含関係がある無限に長いハイアラーキの鎖が存在することを示したが，各正整数kについて$L_n k$（k＝２，３，……）の間に真の包含関係があるかどうかの問題は未解決であり，その解決は困難であると思われる。また，２つの複雑さのクラスが真の包含関係にあるための十分条件を与えたが，この定理で判別されるクラスは，クラスの大きさがありまりにも違いすぎる。したがって，もっと強い結果が望まれる。本論文の言語の複雑さのクラスは，非決定性チューリング機械を標準の機械とし，その認識時間によって定まる複雑さのクラスと関係深いことが予想される。

審 査 結 果 の 要 旨

形式言語理論は，計算機言語のような人工言語の基本的な性質を解明する上に有力な理論である。

最近，言語をその処理に要する過程の複雑さによって類別し，その各クラスの性質や，相互関係を調べるという重要な研究分野が発展しつつある。著者は，句構造言語の複雑さとして，言語の生成過程の複雑さに基づく一つの測度を導入し，この測度によって定義される言語のクラスに関して各クラスの基本的な性質や，異なるクラス間の関係，特にその階層構造などを解明した。本論文はその研究成果をまとめたもので，全編5章よりなる。

第1章では，まず句構造言語について，語の長さと，その語を生成するに要する導出の長さの関係から，言語の複雑さの定義を与えている。ついで，この複雑さに関連してコネクテッド文法の存在，スピードアップの定理など，本研究の基礎となる諸定理を導いている。

第2章は，同一クラスに属する言語間の演算についてそのクラスを保存する演算と，保存しない演算を解明したものである。すなわち，ε ーフリーな準同型写像，和集合，連接，正規集合との共通集合などはクラスを保存する演算であること，ある条件のもとでは，逆準同型写像，a ー変換器による写像，スター演算などもクラスを保存する演算であること，および，ある条件の下では準同型写像，補集合などはクラスを保存しない演算であることを明らかにしている。

第3章は，よく知られた言語のクラスとの関係について述べたものである。すなわち，文脈自由型言語は複雑さの単純なクラスに真に含まれること，文脈規定型言語を含むクラスが存在することなどを明らかにするとともに，抽象化された言語の族，抽象化された受理機械の族，および，準実時間言語の族との関係などを解明している。この結果は言語間の関係に興味深い知識を加えたものである。

第4章では，まず複雑さのクラスの間に階層構造が存在すること，さらに，複雑さのクラスと，抽象化された言語のクラスとの間にも階層構造が存在することを明らかにしている。ついで，句構造言語について，非線形定理が成り立つことを証明し，この結果をもちいてある二つのクラス間に，階層構造をもつ無限に多くのクラスが存在することを示している。最後に，二つのクラスの間に真の包含関係が存在するための十分条件を与えている。これらの結果はきわめて難解な問題を解決したものであって，すぐれた成果である。

第5章は，複雑さのクラスに関する決定問題を論じ，多くの問題が決定不能であることを示したものである。

以上要するに，本論文は言語理論の中心問題の一つである言語の複雑さについて，生成過程の複雑さに基づくクラスを定義し，その各クラスの性質，クラス間の関係などの基本的問題を解明したものであって，情報工学の発展に寄与するところが少なくない。

よって，本論文は工学博士の学位論文として合格と認める。